```
!!NA_SEQUENCE 1.0
LOCUS       HSM800235               1433 bp    mRNA    linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp586B1922 (from clone DKFZp586B1922).
ACCESSION   AL049450
VERSION     AL049450.1  GI:4500236
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1433)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp586B1922) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .1433
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp586B1922"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586B1922"
                     /tissue_type="uterus"
                     /clone_lib="586 (synonym: hute1). Vector pSport1; host
                     DH10B; sites NotI + SalI/MluI"
                     /dev_stage="adult"
     polyA_signal    1386. .1391
     polyA_site      1412
BASE COUNT      414 a    321 c    331 g    367 t
ORIGIN

  AL049450  Length: 1433  July 21, 2003 13:18  Type: N  Check: 450  ..

      1  TAGGAAGAAG CGCTCAACAC AGAACACCCT AAGTATCCCT ACTCCCTTTC
     51  AAGGACTAAA TCCCTTCAGA GATTTCCTTT GCTTCAGTCC TTCAGGAGCC
    101  TCCCGAAGAT GCTAGGAGTC GTTTAAGAAG GCTGGAAGAT GGCCGTTGGG
    151  AAGGAACTGT TGGTAGCTGG CAATACTTCA TCCCAAGGAG GCGGTGGCAA
    201  TAGCTCCCTG GGCAGCAGAC ATCATGGATT TTCCTCCCTC TGGGAACCAG
    251  AGGTCCCTCC GTTTCTGGTT TGAATCCCTT AGAGACAAAG AACTTCATTC
    301  TCTCCCATAC CCGCGTGGGA CTGCTGGGGA GGAGAGGGGG TGGAGGGGTT
    351  AGACTTCTAC CCCCATTTGG AAGTACAGAA AGAAAGACTC CTGTTAAAAA
    401  GCTGGCTCCT GAATTTTCAA AAGTCAAAAC AAAAACTCCT AAGCACTCTC
    451  CTATTAAAGA GGAACCCTGT GGTTCCTTAT CTGAAACTGT TTGTAAACGT
    501  GAATTGAGGA GCCAAGAAAC CCCAGAAAAG CCCCGGTCTT CAGTGGACAC
    551  CCCACCAAGA CTCTCCACTC CCCAAAAGGG ACCCAGCACC CATCCCAAGG
    601  AGAAAGCCTT CTCAAGTGAG ATAGAAGATT TGCCGTACCT TTCCACCACA
    651  GAAATGTATT TGTGTCGTTG GCACCAGCCT CCCCCATCAC CGTTACCATT
    701  ACGGGAATCC TCTCCAAAGA AGGAGGAGAC TGTAGCAAGT AAGGCATAGA
    751  GAACACTTGC TCTTATACCC TAGTGGTGGC GGTCAAGCTA ACAAGTGTGA
    801  AAATGCCTTT GGCATTTTTA AAAAAGTGCA ATCAATAAAG CAGAGTTCTG
    851  TCAAGAATGA GTAAGTTAAC AGCCAGAGAC AGACACTGTG CAGGCATTGC
    901  AAATAGATGG AATTACAGCA AAATGTGCTC AATGTATTTG CCTGCTTACA
    951  ACACTGGGAG ATGTGTTTGC CAGTAAGTTG CTCATCACAA GAGCACCAGA
   1001  CTTGGGGGTG TAATCTCCGG CAACTTGCAT GCCCTCTGAA AGAAGGGTTT
   1051  TCTGTGCTGT GAAATGCATA GAACTATACT TTGCCATGCA CGACTGTTCC
   1101  TGCAATTGAT ATTGTGTGAA ATCTGGGAGG GTGGTCTTTG GGTGTTCTCA
   1151  GGGGCCAATG GTAATTTTTG GGTTGGGGAG CCAGCTTGGG GTGGGGAATT
   1201  TTCACCTGGG CCTCCGCTCT TTAACTATAT AAACATTTAT CTGTATATCT
   1251  ATGTCCCTGT CTGGGGGGCA GGAGGAATCT GCCAAAGACC AACAGTCTTA
   1301  CTTTATCTTA CTATACTTCA CAAAGGTTCT AAAATGTGAA GAGTTTACTT
   1351  GGATTGCAGT AGCCCATTGG TTGTTCATAT ATTTAAATAA AATGGTCTAC
   1401  AAACTATTTT TCAAACAAAA AAAAAAAAAA AAA
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2003, 13:29:30 ; Search time 3 Seconds
(without alignments)
4.177 Million cell updates/sec

Title: HSM800235
Perfect score: 1433
Sequence: 1 TAGGAAGAAGCGCTCAACAC.....AACAAAAAAAAAAAAAAAAA 1433

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 4372 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Database : US10054935.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	380.2	26.5	4372	1	US-10-054-935-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-054-935-1
; Sequence 1, Application US/10054935
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: BREAST CANCER TRANSCRIPTION FACTOR GENE AND USES
; FILE REFERENCE: 16U 107 R1
; CURRENT APPLICATION NUMBER: US/10/054,935
; CURRENT FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1922)
; OTHER INFORMATION:
US-10-054-935-1

Query Match 26.5%; Score 380.2; DB 1; Length 4372;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 391; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy        331 GGAGAGGGGGTGGAGGGGTTAGACTTCTACCCCCATTTGGAAGTACAGAAAGAAAGACTC 390
              || |  | ||    ||   || |     | ||||||||||||||||||||||||||||||
Db       1045 GGCGGAGTGGAAAGGGACATAAAAGGAAATCCCCATTTGGAAGTACAGAAAGAAAGACTC 1104

Qy        391 CTGTTAAAAAGCTGGCTCCTGAATTTTCAAAAGTCAAAACAAAAACTCCTAAGCACTCTC 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1105 CTGTTAAAAAGCTGGCTCCTGAATTTTCAAAAGTCAAAACAAAAACTCCTAAGCACTCTC 1164

Qy        451 CTATTAAAGAGGAACCCTGTGGTTCCTTATCTGAAACTGTTTGTAAACGTGAATTGAGGA 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1165 CTATTAAAGAGGAACCCTGTGGTTCCTTATCTGAAACTGTTTGTAAACGTGAATTGAGGA 1224

Qy        511 GCCAAGAAACCCCAGAAAAGCCCCGGTCTTCAGTGGACACCCCACCAAGACTCTCCACTC 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1225 GCCAAGAAACCCCAGAAAAGCCCCGGTCTTCAGTGGACACCCCACCAAGACTCTCCACTC 1284

Qy        571 CCCAAAAGGGACCCAGCACCCATCCCAAGGAGAAAGCCTTCTCAAGTGAGATAGAAGATT 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1285 CCCAAAAGGGACCCAGCACCCATCCCAAGGAGAAAGCCTTCTCAAGTGAGATAGAAGATT 1344

Qy        631 TGCCGTACCTTTCCACCACAGAAATGTATTTGTGTCGTTGGCACCAGCCTCCCCCATCAC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1345 TGCCGTACCTTTCCACCACAGAAATGTATTTGTGTCGTTGGCACCAGCCTCCCCCATCAC 1404

Qy        691 CGTTACCATTACGGGAATCCTCTCCAAAGAAGGAGGAGACTGTAGCAAG 739
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1405 CGTTACCATTACGGGAATCCTCTCCAAAGAAGGAGGAGACTGTAGCAAG 1453
```

Search completed: July 21, 2003, 13:29:34
Job time : 3 secs

```
!!NA_SEQUENCE 1.0
LOCUS       HSM800235              1433 bp    mRNA    linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp586B1922 (from clone DKFZp586B1922).
ACCESSION   AL049450
VERSION     AL049450.1  GI:4500236
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1433)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp586B1922) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .1433
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp586B1922"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586B1922"
                     /tissue_type="uterus"
                     /clone_lib="586 (synonym: hutel). Vector pSport1; host
                     DH10B; sites NotI + SalI/MluI"
                     /dev_stage="adult"
     polyA_signal    1386. .1391
     polyA_site      1412
BASE COUNT      414 a    321 c    331 g    367 t
ORIGIN

  AL049450  Length: 1433  July 21, 2003 13:18  Type: N  Check: 450  ..

       1  TAGGAAGAAG CGCTCAACAC AGAACACCCT AAGTATCCCT ACTCCCTTTC
      51  AAGGACTAAA TCCCTTCAGA GATTTCCTTT GCTTCAGTCC TTCAGGAGCC
     101  TCCCGAAGAT GCTAGGAGTC GTTTAAGAAG GCTGGAAGAT GGCCGTTGGG
     151  AAGGAACTGT TGGTAGCTGG CAATACTTCA TCCCAAGGAG GCGGTGGCAA
     201  TAGCTCCCTG GGCAGCAGAC ATCATGGATT TTCCTCCCTC TGGGAACCAG
     251  AGGTCCCTCC GTTTCTGGTT TGAATCCCTT AGAGACAAAG AACTTCATTC
     301  TCTCCCATAC CCGCGTGGGA CTGCTGGGGA GGAGAGGGGG TGGAGGGGTT
     351  AGACTTCTAC CCCCATTTGG AAGTACAGAA AGAAAGACTC CTGTTAAAAA
     401  GCTGGCTCCT GAATTTTCAA AAGTCAAAAC AAAAACTCCT AAGCACTCTC
     451  CTATTAAAGA GGAACCCTGT GGTTCCTTAT CTGAAACTGT TTGTAAACGT
     501  GAATTGAGGA GCCAAGAAAC CCCAGAAAAG CCCCGGTCTT CAGTGGACAC
     551  CCCACCAAGA CTCTCCACTC CCCAAAAGGG ACCCAGCACC CATCCCAAGG
     601  AGAAAGCCTT CTCAAGTGAG ATAGAAGATT TGCCGTACCT TTCCACCACA
     651  GAAATGTATT TGTGTCGTTG GCACCAGCCT CCCCCATCAC CGTTACCATT
     701  ACGGGAATCC TCTCCAAAGA AGGAGGAGAC TGTAGCAAGT AAGGCATAGA
     751  GAACACTTGC TCTTATACCC TAGTGGTGGC GGTCAAGCTA ACAAGTGTGA
     801  AAATGCCTTT GGCATTTTTA AAAAAGTGCA ATCAATAAAG CAGAGTTCTG
     851  TCAAGAATGA GTAAGTTAAC AGCCAGAGAC AGACACTGTG CAGGCATTGC
     901  AAATAGATGG AATTACAGCA AAATGTGCTC AATGTATTTG CCTGCTTACA
     951  ACACTGGGAG ATGTGTTTGC CAGTAAGTTG CTCATCACAA GAGCACCAGA
    1001  CTTGGGGGTG TAATCTCCGG CAACTTGCAT GCCCTCTGAA AGAAGGGTTT
    1051  TCTGTGCTGT GAAATGCATA GAACTATACT TTGCCATGCA CGACTGTTCC
    1101  TGCAATTGAT ATTGTGTGAA ATCTGGGAGG GTGGTCTTTG GGTGTTCTCA
    1151  GGGGCCAATG GTAATTTTTG GGTTGGGGAG CCAGCTTGGG GTGGGGAATT
    1201  TTCACCTGGG CCTCCGCTCT TTAACTATAT AAACATTTAT CTGTATATCT
    1251  ATGTCCCTGT CTGGGGGGCA GGAGGAATCT GCCAAAGACC AACAGTCTTA
    1301  CTTTATCTTA CTATACTTCA CAAAGGTTCT AAAATGTGAA GAGTTTACTT
    1351  GGATTGCAGT AGCCCATTGG TTGTTCATAT ATTTAAATAA AATGGTCTAC
    1401  AAACTATTTT TCAAACAAAA AAAAAAAAAA AAA
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2003, 13:29:30 ; Search time 3 Seconds
(without alignments)
4.177 Million cell updates/sec

Title: HSM800235
Perfect score: 1433
Sequence: 1 TAGGAAGAAGCGCTCAACAC.....AACAAAAAAAAAAAAAAAAA 1433

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 4372 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Database : US10054935.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	380.2	26.5	4372	1	US-10-054-935-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-054-935-1
; Sequence 1, Application US/10054935
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: BREAST CANCER TRANSCRIPTION FACTOR GENE AND USES
; FILE REFERENCE: 16U 107 R1
; CURRENT APPLICATION NUMBER: US/10/054,935
; CURRENT FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1922)
; OTHER INFORMATION:
US-10-054-935-1

Query Match 26.5%; Score 380.2; DB 1; Length 4372;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 391; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy        331 GGAGAGGGGGTGGAGGGGTTAGACTTCTACCCCCATTTGGAAGTACAGAAAGAAAGACTC 390
              || |  | ||    ||   || |     | ||||||||||||||||||||||||||||||||

Db       1045 GGCGGAGTGGAAAGGGACATAAAAGGAAATCCCCATTTGGAAGTACAGAAAGAAAGACTC 1104

Qy        391 CTGTTAAAAAGCTGGCTCCTGAATTTTCAAAAGTCAAAACAAAAACTCCTAAGCACTCTC 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1105 CTGTTAAAAAGCTGGCTCCTGAATTTTCAAAAGTCAAAACAAAAACTCCTAAGCACTCTC 1164

Qy        451 CTATTAAAGAGGAACCCTGTGGTTCCTTATCTGAAACTGTTTGTAAACGTGAATTGAGGA 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1165 CTATTAAAGAGGAACCCTGTGGTTCCTTATCTGAAACTGTTTGTAAACGTGAATTGAGGA 1224

Qy        511 GCCAAGAAACCCCAGAAAAGCCCCGGTCTTCAGTGGACACCCCACCAAGACTCTCCACTC 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1225 GCCAAGAAACCCCAGAAAAGCCCCGGTCTTCAGTGGACACCCCACCAAGACTCTCCACTC 1284

Qy        571 CCCAAAAGGGACCCAGCACCCATCCCAAGGAGAAAGCCTTCTCAAGTGAGATAGAAGATT 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1285 CCCAAAAGGGACCCAGCACCCATCCCAAGGAGAAAGCCTTCTCAAGTGAGATAGAAGATT 1344

Qy        631 TGCCGTACCTTTCCACCACAGAAATGTATTTGTGTCGTTGGCACCAGCCTCCCCCATCAC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1345 TGCCGTACCTTTCCACCACAGAAATGTATTTGTGTCGTTGGCACCAGCCTCCCCCATCAC 1404

Qy        691 CGTTACCATTACGGGAATCCTCTCCAAAGAAGGAGGAGACTGTAGCAAG 739
              |||||||||||||||||||||||||||||||||||||||||||||||||

Db       1405 CGTTACCATTACGGGAATCCTCTCCAAAGAAGGAGGAGACTGTAGCAAG 1453
```

Search completed: July 21, 2003, 13:29:34
Job time : 3 secs

```
!!NA_SEQUENCE 1.0
LOCUS       HSM800235              1433 bp    mRNA    linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp586B1922 (from clone DKFZp586B1922).
ACCESSION   AL049450
VERSION     AL049450.1  GI:4500236
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1433)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp586B1922) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .1433
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp586B1922"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586B1922"
                     /tissue_type="uterus"
                     /clone_lib="586 (synonym: hute1). Vector pSport1; host
                     DH10B; sites NotI + SalI/MluI"
                     /dev_stage="adult"
     polyA_signal    1386. .1391
     polyA_site      1412
BASE COUNT      414 a    321 c    331 g    367 t
ORIGIN

  AL049450  Length: 1433  July 21, 2003 13:18  Type: N  Check: 450  ..

      1  TAGGAAGAAG CGCTCAACAC AGAACACCCT AAGTATCCCT ACTCCCTTTC
     51  AAGGACTAAA TCCCTTCAGA GATTTCCTTT GCTTCAGTCC TTCAGGAGCC
    101  TCCCGAAGAT GCTAGGAGTC GTTTAAGAAG GCTGGAAGAT GGCCGTTGGG
    151  AAGGAACTGT TGGTAGCTGG CAATACTTCA TCCCAAGGAG GCGGTGGCAA
    201  TAGCTCCCTG GGCAGCAGAC ATCATGGATT TTCCTCCCTC TGGGAACCAG
    251  AGGTCCCTCC GTTTCTGGTT TGAATCCCTT AGAGACAAAG AACTTCATTC
    301  TCTCCCATAC CCGCGTGGGA CTGCTGGGGA GGAGAGGGGG TGGAGGGGTT
    351  AGACTTCTAC CCCCATTTGG AAGTACAGAA AGAAAGACTC CTGTTAAAAA
    401  GCTGGCTCCT GAATTTTCAA AAGTCAAAAC AAAAACTCCT AAGCACTCTC
    451  CTATTAAAGA GGAACCCTGT GGTTCCTTAT CTGAAACTGT TTGTAAACGT
    501  GAATTGAGGA GCCAAGAAAC CCCAGAAAAG CCCCGGTCTT CAGTGGACAC
    551  CCCACCAAGA CTCTCCACTC CCCAAAAGGG ACCCAGCACC CATCCCAAGG
    601  AGAAAGCCTT CTCAAGTGAG ATAGAAGATT TGCCGTACCT TTCCACCACA
    651  GAAATGTATT TGTGTCGTTG GCACCAGCCT CCCCCATCAC CGTTACCATT
    701  ACGGGAATCC TCTCCAAAGA AGGAGGAGAC TGTAGCAAGT AAGGCATAGA
    751  GAACACTTGC TCTTATACCC TAGTGGTGGC GGTCAAGCTA ACAAGTGTGA
    801  AAATGCCTTT GGCATTTTTA AAAAAGTGCA ATCAATAAAG CAGAGTTCTG
    851  TCAAGAATGA GTAAGTTAAC AGCCAGAGAC AGACACTGTG CAGGCATTGC
    901  AAATAGATGG AATTACAGCA AAATGTGCTC AATGTATTTG CCTGCTTACA
    951  ACACTGGGAG ATGTGTTTGC CAGTAAGTTG CTCATCACAA GAGCACCAGA
   1001  CTTGGGGGTG TAATCTCCGG CAACTTGCAT GCCCTCTGAA AGAAGGGTTT
   1051  TCTGTGCTGT GAAATGCATA GAACTATACT TTGCCATGCA CGACTGTTCC
   1101  TGCAATTGAT ATTGTGTGAA ATCTGGGAGG GTGGTCTTTG GGTGTTCTCA
   1151  GGGGCCAATG GTAATTTTTG GGTTGGGGAG CCAGCTTGGG GTGGGGAATT
   1201  TTCACCTGGG CCTCCGCTCT TTAACTATAT AAACATTTAT CTGTATATCT
   1251  ATGTCCCTGT CTGGGGGGCA GGAGGAATCT GCCAAAGACC AACAGTCTTA
   1301  CTTTATCTTA CTATACTTCA CAAAGGTTCT AAAATGTGAA GAGTTTACTT
   1351  GGATTGCAGT AGCCCATTGG TTGTTCATAT ATTTAAATAA AATGGTCTAC
   1401  AAACTATTTT TCAAACAAAA AAAAAAAAAA AAA
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2003, 13:29:30 ; Search time 3 Seconds
(without alignments)
4.177 Million cell updates/sec

Title: HSM800235
Perfect score: 1433
Sequence: 1 TAGGAAGAAGCGCTCAACAC.....AACAAAAAAAAAAAAAAAAA 1433

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 4372 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Database : US10054935.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	380.2	26.5	4372	1	US-10-054-935-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-054-935-1
; Sequence 1, Application US/10054935
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: BREAST CANCER TRANSCRIPTION FACTOR GENE AND USES
; FILE REFERENCE: 16U 107 R1
; CURRENT APPLICATION NUMBER: US/10/054,935
; CURRENT FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1922)
; OTHER INFORMATION:
US-10-054-935-1

Query Match 26.5%; Score 380.2; DB 1; Length 4372;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 391; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy        331 GGAGAGGGGGTGGAGGGGTTAGACTTCTACCCCCATTTGGAAGTACAGAAAGAAAGACTC 390
              || |  | ||    ||   || |     | ||||||||||||||||||||||||||||||

Db       1045 GGCGGAGTGGAAAGGGACATAAAAGGAAATCCCCATTTGGAAGTACAGAAAGAAAGACTC 1104

Qy        391 CTGTTAAAAAGCTGGCTCCTGAATTTTCAAAAGTCAAAACAAAAACTCCTAAGCACTCTC 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1105 CTGTTAAAAAGCTGGCTCCTGAATTTTCAAAAGTCAAAACAAAAACTCCTAAGCACTCTC 1164

Qy        451 CTATTAAAGAGGAACCCTGTGGTTCCTTATCTGAAACTGTTTGTAAACGTGAATTGAGGA 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1165 CTATTAAAGAGGAACCCTGTGGTTCCTTATCTGAAACTGTTTGTAAACGTGAATTGAGGA 1224

Qy        511 GCCAAGAAACCCCAGAAAAGCCCCGGTCTTCAGTGGACACCCCACCAAGACTCTCCACTC 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1225 GCCAAGAAACCCCAGAAAAGCCCCGGTCTTCAGTGGACACCCCACCAAGACTCTCCACTC 1284

Qy        571 CCCAAAAGGGACCCAGCACCCATCCCAAGGAGAAAGCCTTCTCAAGTGAGATAGAAGATT 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1285 CCCAAAAGGGACCCAGCACCCATCCCAAGGAGAAAGCCTTCTCAAGTGAGATAGAAGATT 1344

Qy        631 TGCCGTACCTTTCCACCACAGAAATGTATTTGTGTCGTTGGCACCAGCCTCCCCCATCAC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1345 TGCCGTACCTTTCCACCACAGAAATGTATTTGTGTCGTTGGCACCAGCCTCCCCCATCAC 1404

Qy        691 CGTTACCATTACGGGAATCCTCTCCAAAGAAGGAGGAGACTGTAGCAAG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1405 CGTTACCATTACGGGAATCCTCTCCAAAGAAGGAGGAGACTGTAGCAAG 1453
```

Search completed: July 21, 2003, 13:29:34
Job time : 3 secs